

RESEARCH ARTICLE

WILEY

# Narrative coherence in multiple forensic interviews with child witnesses alleging physical and sexual abuse

Zsafia A. Szojka<sup>1</sup>  | Annabelle Nicol<sup>2</sup> | David La Rooy<sup>1</sup>

<sup>1</sup>School of Law, Royal Holloway, University of London, London, UK

<sup>2</sup>Department of Psychology, Abertay University, Dundee, UK

**Correspondence**

Zsafia A. Szojka, Royal Holloway, University of London, London, UK.  
Email: szojka.zsafia@gmail.com

**Funding information**

School of Law at Royal Holloway, University of London

## Summary

This study investigated the narrative coherence of children's accounts elicited in multiple forensic interviews. Transcriptions of 56 police interviews with 28 children aged 3–14 years alleging physical and sexual abuse were coded for markers of completeness, consistency and connectedness. We found that multiple interviews increased the completeness of children's testimony, containing on average almost twice as much new information as single interviews, including crucial location, time and abuse-related details. When both contradictions within the same interview and across interviews were considered, contradictions were not more frequent in multiple interviews. The frequency of linguistic markers of connectedness remained stable across interviews. Multiple interviews increase the narrative coherence of children's testimony through increasing their completeness without necessarily introducing contradictions or decreasing causal-temporal connections between details. However, as 'ground truth' is not known in field studies, further investigation of the relationship between the narrative coherence and accuracy of testimonies is required.

## KEYWORDS

child witnesses, forensic interview, multiple interviews

## 1 | INTRODUCTION

The story-telling model of legal decision making (Pennington & Hastie, 1992) suggests that the narrative coherence of witnesses' accounts plays a crucial role in forensic investigations, as coherent narratives, defined in legal terms as stories which are complete, consistent and causally and temporally connected, are more credible than stories lacking in coherence (McAdams, 2006; Pennington & Hastie, 1988; Rideout, 2008). Although the story-telling model was initially conceived to describe the decision-making processes of jurors during a trial, the model has implications for credibility judgments during police investigations too, as cases judged non-credible at this stage never reach a courtroom. Bennett and Feldman (1981) suggest that the narrative coherence of witnesses' stories is especially critical in situations where

credibility judgements need to be made in the absence of corroborating evidence, which is often the case for child abuse investigations.

Children's ability to provide coherent narratives in forensic interviews may be limited by several factors. Firstly, Reese et al.'s (2011) research on the narrative coherence of children's autobiographical accounts suggests that although young children can give truthful and detailed accounts of events that happened to them, the ability to provide chronologically, contextually and thematically coherent narratives does not fully develop until late childhood or adolescence. Constructing a consistent timeline of events may be particularly difficult for young children due to their difficulty with understanding and using temporal concepts (Graffam Walker, Kenniston, Inada, & Caldwell, 2013), especially in cases where they need to give accounts of repeated experiences (Brubacher, Powell, & Roberts, 2014). Despite

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Applied Cognitive Psychology* published by John Wiley & Sons Ltd.

the central role of storytelling in forensic investigations (Pennington & Hastie, 1992) and children's difficulty with recounting their past experiences in a coherent manner (Reese et al., 2011), few studies applied a story-telling framework to evaluate the narrative quality of children's accounts elicited in forensic interviews.

### 1.1 | Models of narrative coherence in developmental psychology

In contrast to the story-telling model's definition of narrative coherence as an umbrella term comprising of the completeness, consistency and connectedness of arguments (McAdams, 2006; Pennington & Hastie, 1992; Rideout, 2008), psychology studies generally use a narrower concept of narrative coherence, broadly defined as a framework for organising event details (Feltis, Powell, & Roberts, 2011). In this form, narrative coherence is described as independent from other measures of the quality of children's stories, including accuracy, consistency and descriptive detail (Brown, Brown, Lewis, & Lamb, 2018; Reese et al., 2011).

Recent laboratory and field studies analysing the narrative coherence of children's accounts of real-life experiences fall into three broad categories. Story-grammar approaches categorise the information provided by the speaker into six grammar elements, including the setting of the story, the initiating event, the protagonist's internal responses, the attempt or action, the consequence of the action and the protagonist's reaction. (Feltis et al., 2011; Feltis, Powell, Snow, & Hughes-Scholes, 2010; Westcott & Kynan, 2004). In contrast, multidimensional models rely on the broader categories of context, chronology, theme and evaluation to measure the narrative coherence of autobiographical memories (Habermas & Bluck, 2000; Habermas and de Silveira (2008); Peterson, Morris, Baker-Ward, & Flynn, 2014; Reese et al., 2011; Morris, Baker-Ward, & Bauer, 2010; Nelson & Fivush, 2004). Finally, the narrative cohesion approach measures the 'degree to which event details are presented in a connected form through temporal and causal relations' (Kulkofsky, Wang, & Ceci, 2008, p. 23).

Whilst multidimensional models, story schema approaches and the narrative cohesion approach are distinct from one another, elements of their conceptualisation of narrative coherence overlap (Brown et al., 2018). For instance, Brown et al. (2018) mapped the multidimensional model's dimensions of chronology, context and evaluation onto the temporal element of setting, the physical and social elements of setting and the reaction and internal response of protagonist, respectively, in the story schema approach. Similarly, temporal markers in the narrative cohesion approach measure the same aspect of recall as the element of temporal setting in the story schema model and the dimension of chronology in multidimensional models.

### 1.2 | Narrative coherence in forensic interviews with children

Initially, the story grammar approach was used to measure the narrative coherence of forensic interviews with child witnesses (Feltis

et al., 2010; Westcott & Kynan, 2004). Analysing Memorandum of Good Practice interviews with children alleging sexual abuse in the United Kingdom, Westcott and Kynan (2004) found that although children's accounts did conform to a basic story structure, the scarcity of details and the lack of a causal-temporal structure in young children's testimonies limited the coherence of their narratives. A third of the cases examined was rated as 'disordered' in terms of causal and temporal relations and the extent of disorder was especially high for pre-schoolers (Westcott & Kynan, 2004). Another aspect of narrative coherence which children of all ages struggled with was describing 'subjective details'; their personal reactions to the abuse. Only 10% of children spontaneously described their physical perceptions and 20% provided descriptions of their emotional state. However, this proportion was increased to 33 and 46.6%, respectively, when interviewers asked questions focused on subjective details. Noting the low proportion of open-ended questions in their sample, Westcott and Kynan (2004) suggested that the use of overly specific questions compromised narrative coherence, especially the temporal and causal organisation of children's accounts.

In addition to the impact of question type on the frequency of story grammar elements in children's testimony (Westcott & Kynan, 2004), laboratory research suggests that children's age and the characteristics of the event they are questioned about also affect the narrative coherence of recall (Feltis et al., 2011; Reese et al., 2011). Relying on the multidimensional model of narrative coherence, Reese et al. (2011) found that even pre-schoolers were able to maintain a topic of conversation, but they found it difficult to establish events in time and place and organise them according to chronology. Similarly, Habermas and de Silveira (2008) found that temporal coherence increased dramatically between the ages of 8 and 12, causal coherence increased most between 12 and 16 years, and thematic coherence was still in development between 16 and 20 years. Using a story grammar approach, Feltis et al. (2011) reported that the frequency of schema elements in children's accounts correlated positively with age. In addition, the children in Feltis et al.'s (2011) sample also recalled more story grammar elements when the delay between the interview and the events children were asked to recall was shorter, and when children experienced repeated occurrences of the same event rather than a single event.

Overall, research on children's ability to provide coherent narratives in forensic interviews suggests that children struggle with providing a causal-temporal framework and describing their subjective reactions to the events they are questioned about (Westcott & Kynan, 2004). Furthermore, children's age and the type of abuse they experienced may limit their ability to form coherent narratives (Feltis et al., 2011; Reese et al., 2011).

### 1.3 | The risks and benefits of multiple forensic interviews with child witnesses

Many children interviewed for legal purposes in the United Kingdom are questioned more than once. Multiple interviews with child witnesses are associated with both risks and benefits with regards to their impact on the quality of the testimonies provided. The existing

literature on the impact of multiple interviews on the quality of children's testimonies was summarised and evaluated by two separate reviews published in 2008 and 2009 (Goodman & Quas, 2008; La Rooy, Lamb, & Pipe, 2009). Both reviews concluded that children's recall can remain highly accurate in multiple interviews when the interviews are conducted using open-ended questions. However, the use of suggestive questions and long delays between interviews were associated with a decrease in the accuracy of children's accounts (Goodman & Quas, 2008; La Rooy et al., 2009).

Recent research supports the conclusion of these reviews in terms of the risks associated with multiple interviews, showing that all kinds of suggestive interviewing methods, including cross-examination techniques (Fogliati & Bussey, 2014; Jack & Zajac, 2014; Righarts, Jack, Zajac, & Hayne, 2015; O'Neill & Zajac, 2013), false memory paradigms (Otgaar, Verschuere, Meijer, & van Oorsouw, 2012), misleading information (London, Bruck, & Melnyk, 2009) and leading questions (Melinder et al., 2010) reduce the accuracy of children's testimony when used in combination with multiple interviews. However, recent studies have also provided further evidence of the potential benefits of multiple interviews, demonstrating that children are able to recall stressful events even after several interviews separated by yearlong delays (Peterson, 2011, 2015) and that under some conditions, providing children with an opportunity to recall events in response to open-ended questions can ameliorate the effects of previous suggestive interviews (Melinder et al., 2010; Righarts et al., 2015).

The benefits of multiple interviews are also supported by a cost-effectiveness analysis showing that the likelihood of convicting a child sexual abuse offender is increased by an estimated 6.1% as a result of multiple interviews, raising the percentage of offenders who are convicted from 22.8 to 28.9% (Block, Foster, Pierce, Berkoff, & Runyan, 2013). The costs associated with multiple interviews in this analysis involved the additional resources required by law enforcement whilst the key benefit was the prevention of further victimisation through identifying, convicting and incarcerating offenders. However, as Block et al. (2013) note, the risks and benefits associated with multiple interviews are contingent upon the quality of the interviews, with additional suggestive or otherwise substandard interviews having few advantages and many risks, including a decrease in children's credibility due to inconsistencies in their reports across interviews, and a potential increase in false conviction rates (Block et al., 2013). Thus, in order to establish the impact of multiple interviews on the outcome of child sexual abuse cases, it is essential to consider the quality of the testimonies elicited, including their completeness, consistency and connectedness.

#### 1.4 | The impact of multiple interviews on the narrative coherence of testimonies

Whilst the accuracy of children's testimonies provided over the course of multiple interviews has been studied extensively, less is known about the impact of multiple interviews on the narrative coherence of testimonies. The completeness and consistency of children's accounts across repeated recall has been the subject of both field research (Hershkowitz &

Termer, 2007; Katz & Hershkowitz, 2012; Leander, 2010; Waterhouse, Ridley, Bull, La Rooy, & Wilcock, 2016) and laboratory studies (e.g., Knutsson, Allwood, & Johansson, 2011; Peterson, 2011, 2015; Price, Connolly, & Gordon, 2016), but the third component of the 'story-telling' model of coherence, the extent to which details are connected through causal and temporal links, is yet to be explored.

Multiple interviews generally increase the completeness of children's testimony as repeated recall occasions allow witnesses to recall new information (La Rooy et al., 2009). In laboratory studies, the completeness of children's recall in a second interview has sometimes exceeded the number of details reported in the first interview (e.g., Knutsson et al., 2011; La Rooy, Pipe, & Murray, 2005). However, even when the number of details reported in each successive interview declines, the overall amount of information reported in the interviews may increase through reminiscence (Erdelyi, 1996). Consistent with laboratory findings indicating that reminiscence leads to an increase in completeness over multiple interviews, several field studies found that a significant amount of new forensically relevant information was recalled in subsequent interviews (Hershkowitz & Termer, 2007; Katz & Hershkowitz, 2012; Leander, 2010; Waterhouse et al., 2016).

As multiple interviews lead to the recall of new information, they inherently decrease the consistency of children's accounts through the addition of new details and the omission of previously mentioned details from later interviews. Examining cross-examination techniques in Scottish courts, Szojka, Andrews, Lamb, Stolzenberg, and Lyon (2017) found that defence lawyers often challenged the credibility of children's statements on the basis of inconsistencies in their accounts, despite laboratory research showing a lack of correlation between the consistency and overall accuracy of witness statements elicited through multiple interviews (Baugerud, Magnussen, & Melinder, 2014; Gilbert & Fisher, 2006). Defence lawyers most commonly referred to contradictions between details mentioned by the witness, rather than additions of new details or omissions of previously mentioned details (Szojka et al., 2017). Previous research on the frequency of contradictions in children's accounts elicited over multiple interviews has been inconclusive; while Katz and Hershkowitz (2012) found no contradictory details within or across interviews, Waterhouse et al. (2016) reported a low number of forensically relevant contradictions. Contrasting results may be partially accounted for by the lower frequency of open-ended questions in Waterhouse et al.'s (2016) sample compared with the NICHD interviews analysed by Katz and Hershkowitz (2012).

Although the connectedness of children's accounts elicited over the course of multiple forensic interviews has not yet been addressed, Habermas and de Silveira's (2008) study analysed the impact of repeated recall on the extent to which events in children's and adults' life narratives were causally and temporally connected. Comparing the thematic, causal and temporal coherence of 8, 12, 16 and 20 year olds in two 15-minute interviews conducted 2 weeks apart, Habermas and de Silveira (2008) found that multiple interviews did not affect the frequency of temporal and causal elements of narrative coherence.

However, the extent to which Habermas and de Silveira's results can be interpreted in a forensic context are limited. First, forensic

interviews are based on a question-answer format, and lengthy uninterrupted narratives are not frequent in interviews with children. The question-answer format might modify the effect of multiple interviews on narrative coherence – it can either serve as a scaffold, as suggested by Bennett and Feldman (1981), or have a disjointing effect on children's accounts, when the questions asked are overly specific (Westcott & Kynan, 2004). Secondly, children are asked to recall events in much more detail during forensic interviews than when questioned about their life stories, which may allow a larger role for reminiscence in later interviews. Third, the youngest children in Habermas and de Silveira's (2008) sample were 8 years old, but the testimonies of younger children may be particularly lacking in causal and temporal connections due to their limited ability to organise their recall according to context and chronology (Reese et al., 2011).

## 1.5 | The present study

Whilst the impact of multiple interviews on the quality of children's recall has been the focus of a wide range of research, only a few studies have analysed real-life forensic testimonies provided over the course of multiple interviews due to the challenges of accessing and analysing such data (Hershkowitz & Terner, 2007; Katz & Hershkowitz, 2012; Leander, 2010; Waterhouse et al., 2016). Unfortunately, the principal measures used in laboratory research to evaluate the quality of children's testimony, such as the number of correct target details or the proportion of accurate details, are not applicable in field research due to researchers' ignorance to 'ground truth'. This study seeks to address the gap in the literature through analysing multiple interviews using a narrative coherence framework to evaluate the quality of children's testimony with forensically relevant measures designed specifically for field research, including the content of children's testimony, the consistency of their recall and the extent of causal-temporal connections in their accounts. To the authors' knowledge, no previous research has examined the narrative coherence of children's accounts across multiple forensic interviews.

The purpose of this study was to investigate the extent to which children construct coherent narratives in multiple forensic interviews and explore the impact of multiple interviews on each of the three components of narrative coherence, as defined by the story-telling model (McAdams, 2006; Pennington & Hastie, 1992; Rideout, 2008); completeness, consistency and connectedness. When conducting multiple interviews, investigators can use the cumulative amount of information recalled in the two interviews rather than only the contents of the second interview, therefore, the overall amount of information available to the interviewers after the two interviews was compared with the amount of information available after the first interview.

Although a new coding scheme was developed to conform more closely to the legal definition of narrative coherence and to allow for the analysis of very long narratives elicited over the course of multiple interviews, the components of narrative coherence in this study can be mapped onto multidimensional models of coherence (Reese

et al., 2011). Specifically, the content types of time and location indicate context, the proportion of sensitive information conveys a measure of theme, subjective details measure evaluative content and chronology is indexed by the frequency of temporal markers.

## 1.6 | Hypotheses

Previous field research strongly suggests that multiple interviews increase the completeness of children's testimonies (Hershkowitz & Terner, 2007; Katz & Hershkowitz, 2012; Leander, 2010; Waterhouse et al., 2016). Therefore, it was expected that (a) multiple interviews will increase the overall number of details, (b) multiple interviews will increase the number of time and location details, (c) multiple interviews will increase the number of subjective details and (d) multiple interviews will increase the number of sensitive details. Based on the results of previous studies investigating the impact of repeated recall on the consistency of witnesses' accounts (Baugerud et al., 2014; Gilbert & Fisher, 2006; Krix et al., 2015), it was expected that children's testimonies will become less consistent as a result of multiple interviews.

Due to the lack of previous research on the connectedness of children's testimonies elicited over multiple forensic interviews, analyses regarding this aspect of narrative coherence were exploratory and no specific hypotheses were established. Scaffolding provided by interviewers and the 'practice effect' associated with repeated recall may facilitate children's recall, leading to a higher frequency of markers of causal and temporal connectedness in multiple interviews. However, it is also possible that children have difficulties with integrating details mentioned over different recall occasions on the same causal-temporal scale. Alternatively, Habermas and de Silveira's (2008) finding that the frequency of causal-temporal connections in children's accounts is not affected by multiple recall occasions may be replicated in forensic interviews.

## 2 | METHOD

### 2.1 | Sample

Transcriptions of 56 forensic interviews with 28 alleged victims of child physical and sexual abuse aged between 3 and 14 years ( $M = 7.83$ ,  $SD = 3$ ), obtained from an existing dataset, were analysed. Due to the challenges associated with accessing data from real-life police investigations, the sample size was small, but comparable to other field studies involving multiple interviews (Leander, 2010:  $N = 27$ ; Waterhouse et al., 2016:  $N = 21$ ). Interviews were conducted and transcribed for the purpose of police investigations. Permission to conduct the study was granted by the ethics committee of the authors' institution ahead of the start of data collection. Children were interviewed between 2 and 11 times, however, only the first and second interviews with each child were examined. The delay between interviews ranged from less than an hour to over a year and a half

( $M = 111.52$  days,  $SD = 159.49$ ). The interviews took place in Scotland ( $N = 17$ ), England ( $N = 6$ ) and the Republic of Ireland ( $N = 5$ ) between 2001 and 2014. The majority of children were interviewed twice by the same interviewer ( $N = 16$ ) whilst some were interviewed by two different interviewers ( $N = 12$ ). The witnesses interviewed included females ( $N = 15$ ) and males ( $N = 13$ ) alleging physical abuse ( $N = 5$ ), sexual abuse ( $N = 17$ ) or both physical and sexual abuse ( $N = 6$ ). Most witnesses described multiple instances of abuse ( $N = 23$ ). Cases involved a single victim ( $N = 12$ ), two co-victims ( $N = 10$ ) or three co-victims ( $N = 6$ ).

## 2.2 | Coding

Only the substantive part of the interviews was analysed, defined as questions asked after the interviewer initially transitioned to probes related to the context of the allegations. The rapport, narrative practice and closure phases of the interview were not analysed. To provide a measure of interview quality, interviewers' utterances were coded according to question type. Codes for children's responses were divided into three main categories; completeness, consistency and connectedness codes. Completeness codes included the number of details in each utterance, their content, their subjectivity and their relevance to the investigations. To measure consistency, each detail was coded according to their novelty and new details were coded as either consistent or inconsistent with previously reported details. Connectedness was coded by identifying local markers of linguistic coherence.

## 2.3 | Question type

The type of question eliciting each detail reported by children was coded according to a modified version of the question type coding guide developed by Lamb, Orbach, Hershkowitz, Esplin, and Horowitz (2007). Examples of question types are presented in Table 1.

1. *Invitations*. Broad open-ended questions encouraging free recall.
2. *Summaries*. Statements summarising details previously mentioned by the child, either verbatim or paraphrased.
3. *Directives*. Open-ended questions encouraging cued recall focused on a topic previously mentioned by the child.

**TABLE 1** Examples of question types

Question type	Example
Invitations	'Tell me everything that happened'
Summaries	'You said it happened in the living room for the first time and then in your bedroom'.
Directives	'You mentioned that he touched you. Where did he touch you?'
Option-posing questions	'Was your mum at home?'
Suggestions	'He touched you, didn't he?'

4. *Option-posing questions*. Closed-ended yes/no or forced choice questions.
5. *Suggestions*. Leading questions and statements referring to details that the child has not mentioned previously.

## 2.4 | Number of details

The number of details, approximately corresponding to the number of clauses, was counted in each of children's utterances. Both independent clauses and subordinate clauses were coded as separate details.

## 2.5 | Content

Children's responses were coded according to the type of information they referred to. Content codes were not exclusive, however, multiple mentions of the same detail type in the same utterance were only coded once. Examples of content codes are presented in Table 2.

1. *Victim*. Mentions of the child himself or herself.
2. *Suspect*. Mentions of the alleged perpetrator of the abuse.
3. *Co-victim*. Mentions of another alleged victim of the same perpetrator.
4. *Witness*. Mentions of other persons.
5. *Location*. Mentions of places or of the positioning of objects or people.
6. *Time*. Mentions of specific dates or times.
7. *Action*. Mentions of actions, including verbal statements.
8. *Object*. Mentions of objects and pets.

## 2.6 | Subjectivity

Details referring to subjective, personal descriptions of the event from the point of view of the victim or other persons were identified and categorised into one of three categories.

1. *Perceptions*. Mentions of visual, auditory or tactile experiences, including pain.

**TABLE 2** Examples of the content codes assigned to children's responses

Example	Content codes
'On Saturday, my dad bought me a new bicycle'.	Time, suspect, victim, action, object
'My mum shouted at Mr P'.	Witness, action, suspect
'The living room had a large sofa since last Christmas'.	Location, object, time

2. *Emotions*. Mentions of affective processes, such as embarrassment, happiness or anger. Expressions of preferences ('I liked/did not like it'.) were included in this category.
3. *Cognitions*. Mentions of thoughts, aims and beliefs.

## 2.7 | Relevance to the investigation

Children's responses were coded according to whether they were sensitive, central or peripheral from the point of view of the investigators. *Sensitive* content was defined as details describing the time and place of the abuse and details related to the immediate lead up to the abuse, the content of the abuse, or the immediate aftermath of the abuse. *Central* content was defined as details directly related to the location of the abuse, the time scale of the abuse, the alleged perpetrator of the abuse, co-victims and witnesses of the abuse and essential family relationships without referring to specific instances of abuse. *Peripheral* content was defined as focused on details that were only indirectly related to the events that the interview focused on, such as non-plot related descriptions of time, place, people or events.

## 2.8 | Novelty

Each detail mentioned by children was coded according to whether they were *new* or *repeated*. Repeated details were also categorised according to whether they were *repeated within the same interview* or *repeated across interviews*. Details could be coded as repeated both within the same interview and across interviews.

## 2.9 | Consistency

Each repeated detail was coded either as *consistent* or *inconsistent* with previous mentions of the same detail. Inconsistent details were also coded on the basis of whether they contradicted details mentioned *within the same interview* or *across interviews*. Details could be coded as contradictory both with details within the same interview and across interviews. If a detail was once mentioned inconsistently, all future mentions were coded as inconsistent, unless an explanation was provided to resolve the contradiction.

## 2.10 | Connectedness

Local markers of linguistic coherence were identified using a modified version of the coding scheme developed by Kulkofsky et al. (2008). Marker types were not exclusive, however, multiple mentions of the same marker type in the same utterance were only coded once. Examples of marker types are presented in Table 3.

1. *Simple temporal markers*. Phrases signifying chronological order.

**TABLE 3** Examples of linguistic markers of connectedness

Type of marker	Examples
Simple temporal markers	First, next, then, before, after, etc.
Complex temporal markers	When, until, while, etc.
Markers of causal relations	Because, so, in order to, as, etc.
Markers of optional states	Sometimes, usually, always, probably, etc.

2. *Complex temporal markers*. Phrases placing an event in time with relation to other events.
3. *Markers of causal relations*. Phrases describing cause-and-effect relationships between details.
4. *Markers of optional states*. Phrases referring to conditional events.

## 2.11 | Inter-rater reliability

A random selection of 20% of the transcripts ( $N = 12$ ) were independently re-coded. Cohen's  $\kappa$  was used to assess agreement between coders. Reliability was very good for question type,  $\kappa = 0.90$ ,  $SE = 0.01$ , 95% CI [0.88, 0.92], relevance  $\kappa = 0.89$ ,  $SE = 0.01$ , 95% CI [0.87, 0.91], novelty  $\kappa = 0.92$ ,  $SE = 0.02$ , 95% CI [0.88, 0.96], consistency,  $\kappa = 0.91$ ,  $SE = 0.02$ , 95% CI [0.87, 0.94] and markers of connectedness,  $\kappa = 0.95$ ,  $SE = 0.01$ , 95% CI [0.93, 0.97]. Reliability was substantial for content,  $\kappa = 0.77$ ,  $SE = 0.04$ , 95% CI [0.69, 0.85] and subjective details  $\kappa = 0.77$ ,  $SE = 0.04$ , 95% CI [0.69, 0.85].

## 3 | RESULTS

### 3.1 | Preliminary analyses and excluded variables

Initial analyses using mixed ANOVAS showed that measures of the completeness, consistency and coherence of children's testimony (the number of details reported, the consistency of details, the relevance of details for the investigation, the content of details and the frequency of markers of connectedness) were not affected by the gender of the witness, the identity of the interviewer, the delay between interviews, the frequency of the abuse and the type of abuse reported. Therefore, these variables were excluded from further analyses.

Children's age was normally distributed (Kolmogorov-Smirnov:  $D[28] = 0.13$ ,  $p = .20$ ) and included in the analyses as a continuous variable. Children's age was added as a covariate to analyses where previous research suggested a relationship between age and the dependent variable, and the scatterplot also suggested a linear relationship. Although ANCOVAs are often used by researchers to statistically control for the effects a confounding variable on the dependent variable (Schneider, Avivi-Reich, & Mozuraitis, 2015), in the current study, the aim was to assess potential interactions between the effects of the covariate and the dependent variable. Therefore, an interaction term between age and the dependent variables was



included in each analysis. In within-subjects designs, the covariate needs to be centered by subtracting the mean of the covariate from each covariate value to avoid an increase in Type 1 error rates or a loss of power (Schneider et al., 2015). Thus, children's age was centered when used as a covariate in within-subjects analyses. As the probability of Type 1 errors is elevated in within-subject designs involving an interaction between the covariate and the dependent variable (Schneider et al., 2015), alternative statistical tests were conducted following each ANCOVA, without the covariate. These confirmed the main effect in each analysis (Appendix).

### 3.2 | Interview quality

Although interviewers' questions were not the focus of the present study, the type of question eliciting each detail was coded to provide a measure of interview quality. Most details were elicited using directives, followed by option-posing questions, invitations, suggestions and summaries (Table 4). In contrast to best practice guidelines, 7.6% ( $SD = 12.8\%$ ) of details in the first interview and 10.6% ( $SD = 14.7\%$ ) in the second interview were provided in response to suggestions.

To determine whether significantly more details were elicited by some questions types than others and whether the proportion of details elicited by each question type differed in the first and second interview, a RM-ANOVA was conducted assessing the potential effect of question type (invitations, summaries, directives, option-posing questions, suggestions) and interview number (first interview, second interview) on the number of details reported by children. Mauchley's test indicated that the assumption of sphericity was violated for question type,  $\chi^2(9) = 22.54$ ,  $p = .007$ . Consequently, a Greenhouse–Geisser correction was applied ( $\epsilon = .73$ ). A significant main effect was found for question type,  $F(2.93, 82.109) = 33.46$ ,  $p < .001$ ,  $\eta_p^2 = 0.55$ . There was no significant main effect for interview number,  $F(1, 108) = 0.01$ ,  $p = .95$ ,  $\eta_p^2 < 0.001$ . There was a significant interaction between question type and interview number,  $F(4, 108) = 3.41$ ,  $p = .01$ ,  $\eta_p^2 = 0.11$ .

Pairwise comparisons using the Bonferroni adjustment (adjusted alpha levels  $p < .005$ ) for multiple comparisons revealed that overall, significantly more details were elicited by directives and option-posing questions than by invitations, summaries or suggestions. More details

were elicited using invitations than summaries. All pairwise comparisons are presented in Table 5.

The two-way interaction between question type and interview number was followed up with 25 paired samples  $t$  tests comparing the proportion of details elicited by each question type in the first interview and the second interview (adjusted alpha level  $p < .002$ ). In the first interview, a higher proportion of details were elicited using directives and option-posing questions than summaries and suggestions. Invitations elicited a lower proportion of details than directives, but a higher proportion than summaries. In the second interview, directives and option-posing questions elicited more details than invitations, summaries and suggestions. There was no difference between the first interview and second interview in the proportion of details elicited by any of the question types. No other comparisons were significant.

### 3.3 | Completeness

On average, witnesses reported 121.75 ( $SD = 147.49$ ) new details in the first interview and 99.29 ( $SD = 123.14$ ) new details in the second. Table 6 contains the number of details children reported in each content category and the proportion of those details from all new details (the same details could belong to multiple content categories). The number of details reported correlated positively with children's age in both Interview 1 and Interview 2 (Figure 1).

### 3.4 | Number of details

To assess whether significantly more details were reported over the course of the two interviews than in the first one only, a RM-ANCOVA was conducted assessing the potential effect of interview number (first interview, overall) on the number of details witnesses reported. Children's age was added to the analysis as a covariate. Significant main effects were found for interview number,  $F(1, 26) = 20.44$ ,  $p < .001$ ,  $\eta_p^2 = 0.44$ , and children's age,  $F(1, 26) = 11.74$ ,  $p = .002$ ,  $\eta_p^2 = 0.31$ . There was a significant interaction between interview number and children's age,  $F(1, 26) = 4.32$ ,  $p = .048$ ,  $\eta_p^2 = 0.14$ . Inspection of Figure 1 suggests that the difference between the number of details reported in the first interview and over the two interviews increased as the age of witnesses increased.

### 3.5 | Content

Most of the new details mentioned focused on the victim, the suspect, actions, witnesses and locations. Mentions of objects, co-victims and time were less frequent in both interviews, as were subjective details, including emotions, perceptions and cognitions. Children reported less new details in each content category in the second interview, but the proportional frequency of each type of detail remained similar.

**TABLE 4** The proportion of details elicited by each question type in Interview 1 and Interview 2

	Interview 1		Interview 2	
	M	SD	M	SD
Invitations	0.20	0.20	0.13	0.16
Summaries	0.03	0.05	0.08	0.10
Directives	0.40	0.15	0.35	0.14
Option-posing	0.29	0.14	0.33	0.13
Suggestions	0.08	0.13	0.11	0.15

**TABLE 5** Summary of analyses of the proportion of details elicited by each question type in Interview 1 and Interview 2

Analysis			Results			
Type	Variables	Df	Test value	p	Effect size	95% CI
Main effect	Interview	1,27	$F = 0.01$	$=.05$	$\eta_p^2 < 0.001$	
Main effect	Question type	2,93,79.23	$F = 33.46$	$<.001^*$	$\eta_p^2 = 0.55$	
Interaction	Interview X question type	4,108	$F = 3.41$	$=.01^*$	$\eta_p^2 = 0.11$	
Follow up	Invitation vs. summary	27	$t = 3.54$	$=.001^{**}$	$d = 0.94$	[0.09, 0.35]
Follow up	Invitation vs. directive	27	$t = -5.13$	$<.001^{**}$	$d = 1.55$	[-0.59, -0.25]
Follow up	Invitation vs. option-posing	27	$t = -3.46$	$=.002^{**}$	$d = 1.06$	[-0.47, -0.12]
Follow up	Invitation vs. suggestive	27	$t = 1.72$	$=.10$	$d = 0.61$	[-0.03, 0.33]
Follow up	Summary vs. directive	27	$t = -10.26$	$<.001^{**}$	$d = 3.26$	[-0.77, -0.51]
Follow up	Summary vs. option-posing	27	$t = -10.33$	$<.001^{**}$	$d = 2.71$	[-0.62, -0.41]
Follow up	Summary vs. suggestive	27	$t = -1.49$	$.15$	$d = 0.33$	[-0.17, 0.03]
Follow up	Directive vs. option-posing	27	$t = 1.81$	$.08$	$d = 0.61$	[-0.12, 0.27]
Follow up	Directive vs. suggestive	27	$t = 8.92$	$<.001^{**}$	$d = 2.52$	[0.44, 0.70]
Follow up	Option-posing vs. suggestive	27	$t = 7.62$	$<.001^{**}$	$d = 2.00$	[0.33, 0.56]
Follow up	Invitation 1 vs. summary 1	27	$t = 4.18$	$<.001^{***}$	$d = 1.23$	[0.09, 0.26]
Follow up	Invitation 1 vs. directive 1	27	$t = -3.47$	$=.001^{***}$	$d = 1.07$	[-0.31, -0.08]
Follow up	Invitation 1 vs. option-posing 1	27	$t = -1.40$	$=.17$	$d = 0.46$	[-0.20, 0.04]
Follow up	Invitation 1 vs. suggestive 1	27	$t = 3.63$	$=.01$	$d = 0.82$	[0.03, 0.23]
Follow up	Summary 1 vs. directive 1	27	$t = -12.38$	$<.001^{***}$	$d = 3.31$	[-0.43, -0.31]
Follow up	Summary 1 vs. option-posing 1	27	$t = -9.70$	$<.001^{***}$	$d = 2.47$	[-0.31, -0.20]
Follow up	Summary 1 vs. suggestive 1	27	$t = -2.38$	$=.03$	$d = 0.63$	[-0.08, -0.01]
Follow up	Directive 1 vs. option-posing 1	27	$t = 2.55$	$=.02$	$d = 0.76$	[0.02, 0.21]
Follow up	Directive 1 vs. suggestive 1	27	$t = 8.84$	$<.001^{***}$	$d = 2.51$	[0.25, 0.40]
Follow up	Option-posing 1 vs. suggestive 1	27	$t = 7.17$	$<.001^{***}$	$d = 1.73$	[0.15, 0.27]
Follow up	Invitation 2 vs. summary 2	27	$t = 1.07$	$=.29$	$d = 0.27$	[-0.04, 0.13]
Follow up	Invitation 2 vs. directive 2	27	$t = -5.41$	$<.001^{***}$	$d = 1.53$	[-0.31, -0.14]
Follow up	Invitation 2 vs. option-posing 2	27	$t = -4.73$	$<.001^{***}$	$d = 1.51$	[-0.31, -0.12]
Follow up	Invitation 2 vs. suggestive 2	27	$t = -0.41$	$.69$	$d = 0.06$	[-0.08, 0.12]
Follow up	Summary 2 vs. directive 2	27	$t = -6.22$	$<.001^{***}$	$d = 2.00$	[-0.36, -0.18]
Follow up	Summary 2 vs. option-posing 2	27	$t = -7.87$	$<.001^{***}$	$d = 2.00$	[-0.33, -0.19]
Follow up	Summary 2 vs. suggestive 2	27	$t = -0.67$	$.51$	$d = 0.21$	[-0.11, 0.06]
Follow up	Directive 2 vs. option-posing 2	27	$t = 0.29$	$.78$	$d = 0.07$	[-0.07, 0.09]
Follow up	Directive 2 vs. suggestive 2	27	$t = 5.99$	$<.001^{***}$	$d = 1.65$	[0.16, 0.33]
Follow up	Option-posing 2 vs. suggestive 2	27	$t = 5.44$	$<.001^{***}$	$d = 1.64$	[0.15, 0.32]
Follow up	Invitation 1 vs. invitation 2	27	$t = 2.21$	$=.04$	$d = 0.50$	[0.01, 0.16]
Follow up	Summary 1 vs. summary 2	27	$t = -1.87$	$=.07$	$d = 0.51$	[-0.10, 0.004]
Follow up	Directive 1 vs. directive 2	27	$t = 1.60$	$=.12$	$d = 0.34$	[-0.01, 0.11]
Follow up	Option-posing 1 vs. option-posing 2	27	$t = -1.78$	$=.08$	$d = 0.37$	[-0.11, 0.01]
Follow up	Suggestive 1 vs. suggestive 2	27	$t = -1.24$	$=.23$	$d = 0.15$	[-0.08, 0.02]

Note: \*Denotes significance at  $p < .05$ , \*\* denotes significance at adjusted alpha level  $p < .005$ , \*\*\* denotes significance at adjusted alpha level  $p < .002$ .

To assess whether significantly more details were reported in each content category over the course of the two interviews than in the first one only, a two-way RM ANOVA was conducted investigating the potential effect of interview number (first interview, overall) on the number of details reported in each content category (suspect,

victim, witness, co-victim, action, object, location, time). Mauchley's test indicated that the assumption of sphericity was violated for content,  $\chi^2(27) = 378.95$ ,  $p < .001$ , and the interaction between content and interview number,  $\chi^2(27) = 323.33$ ,  $p < .001$ . Consequently, Greenhouse–Geisser corrections were applied ( $\epsilon = .20$ ,  $\epsilon = .23$ ,



respectively). Significant main effects were found for interview number,  $F(1, 27) = 18.17, p < .001, \eta_p^2 = 0.40$  and content,  $F(1.39, 37.39) = 23.45, p < .001, \eta_p^2 = 0.40$ . There was a significant

**TABLE 6** The number of details reported in each content category in Interview 1 and Interview 2

	Interview 1		Interview 2	
	M	SD	M	SD
New details	<b>121.75</b>	<b>147.49</b>	<b>99.29</b>	<b>123.14</b>
Victim	49.04	57.86	46.32	53.76
Suspect	45.39	52.11	37.11	40.52
Witness	26	42.3	23.47	45.9
Co-victim	12.11	18.88	6.68	11.26
Action	62.07	80.13	54.86	74.62
Object	13.14	19.31	11.43	18.57
Location	24.14	41.02	15.57	19.23
Time	9.14	16.68	7.57	10.19
Subjectivity	<b>16.71</b>	<b>20.07</b>	<b>16.71</b>	<b>20.07</b>
Cognition	2.75	3.79	2.71	4.08
Perception	3.82	5.14	4.57	6.44
Emotion	10.14	11.14	9.43	13.62

Note: The bold values represent the total number of new details and subjective details in all subcategories. The number of details in each content category does not add up to the total number of new details, as a single detail could belong to multiple content categories.

interaction between interview number and content,  $F(1.62, 43.68) = 12.77, p < .001, \eta_p^2 = 0.32$ .

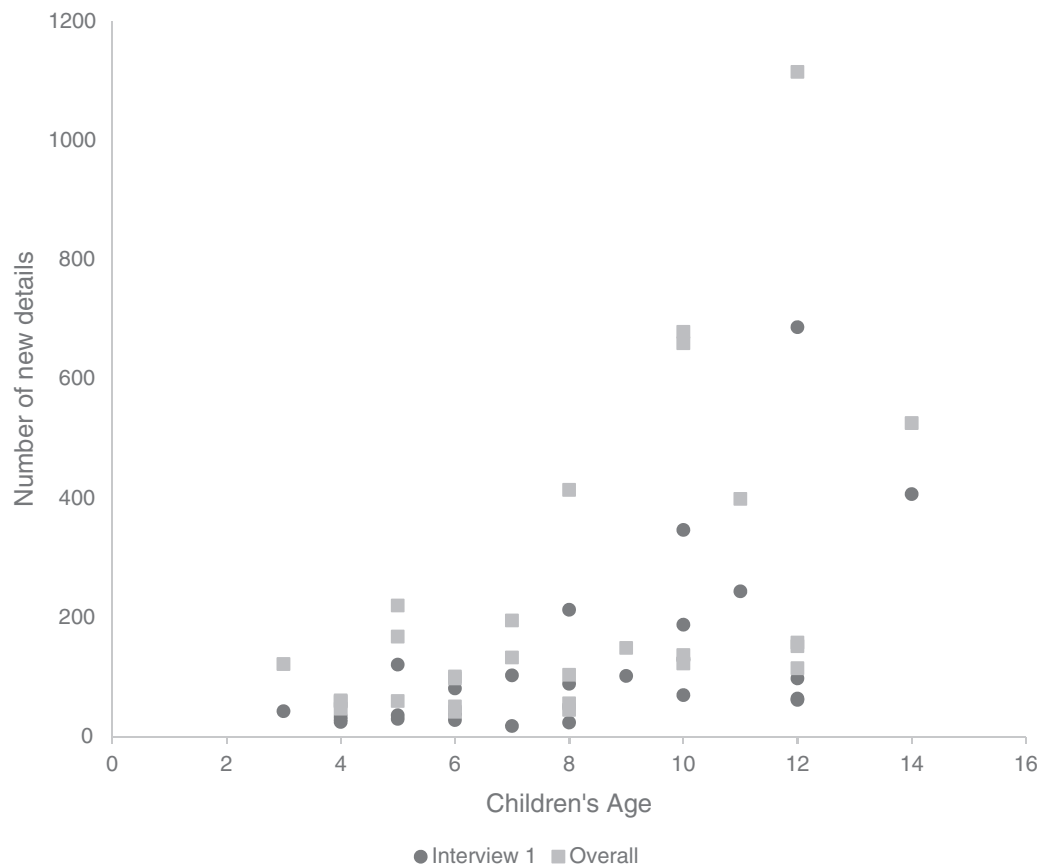
Pairwise comparisons using the Bonferroni adjustment (adjusted alpha levels  $p < .001$ ) for multiple comparisons revealed that significantly more details were reported related to actions, the victim and the suspect than related to objects, witnesses, co-victims, locations and time. Significantly more location-related details were reported than temporal details. All pairwise comparisons are presented in Table 7.

The two-way interaction between interview number and content was followed up with eight paired samples *t* tests comparing the number of details reported in each category in Interview 1 and overall (adjusted alpha level  $p < .006$ ). Significantly more details were reported overall than in the first interview only in all content categories with the exception of information related to witnesses.

### 3.6 | Subjectivity

On average, witnesses reported 16.71 new subjective details in both the first interview and the second interview (Table 3). In both interviews, emotions were the most frequently mentioned subjective details, followed by perceptions and cognitions.

To assess whether significantly more subjective details were reported altogether in the two interviews than from the first interview only, a two-way RM ANOVA was conducted investigating the potential effect of interview number (first interview, overall) on the number



**FIGURE 1** Children's age and the number of new details in Interview 1 and overall

**TABLE 7** Summary of analyses of the content of children's responses in Interview 1 and overall

Analysis			Results			
Type	Variables	Df	Test value	p	Effect size	95% CI
Main effect	Interview	1,26	$F = 18.17$	$<.001^*$	$\eta_p^2 = 0.40$	
Main effect	Content	1.38,37.39	$F = 18.16$	$<.001^*$	$\eta_p^2 = 0.40$	
Interaction	Interview X content	1.62,43.68	$F = 12.77$	$<.001^*$	$\eta_p^2 = 0.32$	
Follow up	Action vs. object	27	$t = 3.98$	$=.012$	$d = 0.89$	[44.68, 140.03]
Follow up	Action vs. victim	27	$t = 2.09$	$=.05$	$d = 0.17$	[.44,42.71]
Follow up	Action vs. suspect	27	$t = 2.60$	$=.02$	$d = 0.30$	[7.23, 61.63]
Follow up	Action vs. witness	27	$t = 5.45$	$<.001^{**}$	$d = 0.62$	[42.06, 92.87]
Follow up	Action vs. co-victim	27	$t = 4.23$	$<.001^{**}$	$d = 0.95$	[50.52, 145.76]
Follow up	Action vs. location	27	$t = 4.10$	$<.001^{**}$	$d = 0.70$	[38.11, 114.31]
Follow up	Action vs. time	27	$t = 4.35$	$<.001^{**}$	$d = 0.97$	[52.95, 147.82]
Follow up	Victim vs. object	27	$t = 4.91$	$<.001^{**}$	$d = 0.94$	[41.18, 100.39]
Follow up	Suspect vs. object	27	$t = 5.40$	$<.001^{**}$	$d = 0.94$	[35.91, 79.95]
Follow up	Witness vs. object	27	$t = 1.80$	$=.08$	$d = 0.38$	[−3.47,53.26]
Follow up	Co-victim vs. object	27	$t = -0.88$	$=.39$	$d = 0.19$	[−19.28, 7.71]
Follow up	Object vs. location	27	$t = -1.98$	$=.06$	$d = 0.35$	[−32.84, 0.56]
Follow up	Object vs. time	27	$t = 1.68$	$=.11$	$d = 0.26$	[−1.74, 17.46]
Follow up	Suspect vs. victim	27	$t = -2.22$	$=.04$	$d = 0.10$	[−24.73, −0.98]
Follow up	Victim vs. witness	27	$t = 5.70$	$<.001^{**}$	$d = 0.49$	[29.38, 62.40]
Follow up	Victim vs. co-victim	27	$t = 4.84$	$<.001^{**}$	$d = 1.04$	[44.13, 109.01]
Follow up	Victim vs. location	27	$t = 5.24$	$<.001^{**}$	$d = 0.67$	[33.26, 76.03]
Follow up	Victim vs. time	27	$t = 5.27$	$<.001^{**}$	$d = 1.07$	[48.05, 109.24]
Follow up	Suspect vs. witness	27	$t = 4.00$	$<.001^{**}$	$d = 0.39$	[16.07, 50.00]
Follow up	Suspect vs. co-victim	27	$t = 5.30$	$<.001^{**}$	$d = 1.06$	[39.07, 88.36]
Follow up	Suspect vs. location	27	$t = 4.90$	$<.001^{**}$	$d = 0.60$	[24.30, 59.27]
Follow up	Suspect vs. time	27	$t = 5.88$	$<.001^{**}$	$d = 1.10$	[42.84, 88.73]
Follow up	Witness vs. co-victim	27	$t = 2.49$	$=.02$	$d = 0.48$	[5.47, 55.89]
Follow up	Witness vs. location	27	$t = 1.13$	$=.27$	$d = 0.12$	[−7.09, 24.59]
Follow up	Witness vs. time	27	$t = 2.65$	$=.01$	$d = 0.52$	[7.38, 58.12]
Follow up	Co-victim vs. location	27	$t = -3.12$	$=.004$	$d = 0.50$	[−36.35, −7.51]
Follow up	Co-victim vs. time	27	$t = 0.65$	$=.52$	$d = 0.08$	[−4.46, 8.60]
Follow up	Location vs. time	27	$t = 3.78$	$<.001^{**}$	$d = 0.55$	[10.97, 37.03]
Follow up	Action 1 vs. action all	27	$t = 3.89$	$=.001^{***}$	$d = 0.47$	[25.92, 83.79]
Follow up	Object 1 vs. object all	27	$t = 3.26$	$=.003^{***}$	$d = 0.45$	[4.23, 18.63]
Follow up	Victim 1 vs. victim all	27	$t = 4.56$	$<.001^{***}$	$d = 0.56$	[25.47, 67.17]
Follow up	Suspect 1 vs. suspect all	27	$t = 4.85$	$<.001^{***}$	$d = 0.55$	[21.40, 52.82]
Follow up	Witness 1 vs. witness all	27	$t = 2.71$	$=.01$	$d = 0.35$	[5.66, 41.26]
Follow up	Co-victim 1 vs. co-victim all	27	$t = 3.14$	$=.004^{***}$	$d = 0.29$	[2.31, 11.04]
Follow up	Location 1 vs. location all	27	$t = 4.56$	$<.001^{***}$	$d = 0.34$	[9.11, 24.03]
Follow up	Time 1 vs. time all	27	$t = 3.93$	$=.001^{***}$	$d = 0.35$	[3.62, 11.52]

Note: \* denotes significance at  $p < .05$ , \*\* denotes significance at adjusted alpha level  $p < .001$ , \*\*\* denotes significance at adjusted alpha level  $p < .006$ .

of details reported in each subjective content category (emotion, cognition, perception). Mauchley's test indicated that the assumption of sphericity was violated for subjectivity,  $\chi^2(2) = 31.57$ ,  $p < .001$ , and the interaction between subjectivity and interview number,  $\chi^2(2) = 20.6$ ,  $p < .001$ . Consequently, Greenhouse–Geisser corrections

were applied ( $\epsilon = .59$ ,  $\epsilon = .65$ , respectively). Significant main effects were found for interview number,  $F(1, 27) = 15.84$ ,  $p < .001$ ,  $\eta_p^2 = 0.37$  and subjectivity,  $F(1.17, 31.71) = 12.70$ ,  $p = .001$ ,  $\eta_p^2 = 0.32$ . There was a significant interaction between interview number and subjectivity,  $F(1.29, 34.9) = 8.51$ ,  $p = .003$ ,  $\eta_p^2 = 0.24$ .

Pairwise comparisons using the Bonferroni adjustment (adjusted alpha levels  $p < .017$ ) for multiple comparisons revealed that significantly more subjective details focused on emotions than on perceptions or cognitions. All pairwise comparisons are presented in Table 8.

The two-way interaction between interview number and subjectivity was followed up with three paired samples  $t$  tests comparing the number of details reported in each subjective content category in Interview 1 and overall (adjusted alpha level  $p < .017$ ). Significantly more details were reported overall than in the first interview about emotions, perceptions and cognitions.

### 3.7 | Sensitive details

Child witnesses reported a larger number of new sensitive details in the second interview ( $M = 41.57$ ,  $SD = 53.58$ ) than in the first interview ( $M = 32.07$ ,  $SD = 48.69$ ). The percentage of sensitive details from all details witnesses reported increased from 30.6% ( $SD = 25\%$ ) in the first interview to 48.1% ( $SD = 29\%$ ) in the second interview. A paired  $t$  test showed that interviewers were aware of significantly more new sensitive details after the second interview than after the first interview,  $t(27) = 4.09$ ,  $p < .001$ ,  $d = 0.63$ , 95% CI = [20.69, 62.45]. A second paired  $t$  test showed that the increase in the percentage of sensitive details from the first to the second interview was significant,  $t(27) = 2.49$ ,  $p = .02$ ,  $d = 0.62$ , 95% CI = [0.03, 0.32].

### 3.8 | Consistency

In the first interview, 23% ( $SD = 12\%$ ) of all information reported was repeated, whilst in the second interview this percentage increased to 38.5% ( $SD = 15\%$ ). Overall, 32% ( $SD = 0.12\%$ ) of details reported in the two interviews were repeated. In the second interview, information was more frequently repeated within the interview ( $M = 0.28$ ,  $SD = 0.12$ ) than across the two interviews ( $M = 0.17$ ,  $SD = 0.16$ ). Some

information was repeated both within the same interview and across the two interviews.

The percentage of consistent repeated details decreased from 87.1% ( $SD = 15\%$ ) in Interview 1 to 84.0% ( $SD = 16\%$ ) in Interview 2. Considering details reported over the course of the two interviews, 84.1% ( $SD = 16\%$ ) of repeated details were consistent with previously reported information. Accordingly, inconsistencies within the same interview ( $M = 0.08$ ,  $SD = 0.12$ ) were also more common than inconsistencies across interviews ( $M = 0.04$ ,  $SD = 0.06$ ). Some repeated information was inconsistent with previously reported information in the same interview as well as in the previous interview ( $M = 0.04$ ,  $SD = 0.06$ ). Older children reported a higher percentage of consistent information than younger children in the first interview, although not overall (Figure 2).

To assess whether the percentage of consistent responses was affected by interview number (first interview, overall), a RM ANCOVA was conducted investigating the potential effect of interview number (first interview, overall) on the percentage of consistent responses. Children's age was included as a covariate in the analysis. One witness provided no repeated information in either interview, therefore this witness was excluded from the analysis. The effect of interview number on consistency was non-significant,  $F(1,26) = 0.47$ ,  $p = .50$ ,  $\eta_p^2 = 0.02$ . There was no significant effect of age,  $F(1,26) = 2.50$ ,  $p = .13$ ,  $\eta_p^2 = 0.09$ . There were no significant interactions.

### 3.9 | Markers of connectedness

In the first interview, 15.6% ( $SD = 9\%$ ) of details mentioned by children contained markers of connectedness whilst this number increased to 17.9% ( $SD = 9\%$ ) in the second interview. The frequency of simple temporal markers, complex temporal markers and causal connections increased from the first to the second interview whilst the frequency of markers of optional states decreased (Table 9). Markers of connectedness were more frequently used by older witnesses than younger witnesses (Figure 3).

**TABLE 8** Summary of analyses of the subjective content of children's responses in Interview 1 and overall

Analysis			Results			
Type	Variables	Df	Test value	p	Effect size	95% CI
Main effect	Interview	1,27	$F = 15.84$	$<.001^*$	$\eta_p^2 = 0.37$	
Main effect	Subjectivity	1,17,31,71	$F = 12.70$	$=.001^*$	$\eta_p^2 = 0.32$	
Interaction	Interview X subjectivity	1,29,34,9	$F = 8.51$	$=.003^*$	$\eta_p^2 = 0.24$	
Follow up	Cognition vs. perception	27	$t = -2.27$	$=.03$	$d = 0.35$	[-5.57, -.28]
Follow up	Cognition vs. emotion	27	$t = -3.77$	$=.001^{**}$	$d = 0.83$	[-21.77, -6.44]
Follow up	Perception vs. emotion	27	$t = -4.44$	$=.002^{**}$	$d = 0.63$	[-17.91, -4.45]
Follow up	Emotion 1 vs. emotion all	27	$t = -3.66$	$=.001^{**}$	$d = 0.52$	[-14.71, -4.15]
Follow up	Cognition 1 vs. cognition all	27	$t = -3.52$	$=.002^{**}$	$d = 0.49$	[-4.3, -1.13]
Follow up	Perception 1 vs. perception all	27	$t = -2.07$	$=.001^{**}$	$d = 0.60$	[-7.07, -2.07]

Note: \*Denotes significance at  $p < .05$ , \*\* denotes significance at adjusted alpha level  $p < .017$ .

To assess whether children's accounts were significantly more connected in the second interview than in the first one, a RM-ANCOVA was conducted assessing the potential effect of interview number (first interview, overall) on the frequency of the four types of connectedness markers (simple temporal, complex temporal, causal, optional) witnesses reported. Children's age was added to the analysis as a covariate. Significant main effects were found for connectedness,  $F(3, 78) = 4.26$ ,  $p = .008$ ,  $\eta_p^2 = 0.14$ , and children's age,  $F(1, 26) = 25.02$ ,  $p < .001$ ,  $\eta_p^2 = 0.49$ . No main effect was found for interview number,  $F(1, 26) = 1.41$ ,  $p = .25$ ,  $\eta_p^2 = 0.05$ . No significant interactions were found.

To follow up the main effect of connectedness, six pairwise comparisons were conducted on the types of connectedness markers using the Bonferroni adjustment (adjusted alpha levels  $p < .008$ ) for multiple comparisons. Markers of causal connections were significantly more frequently used than markers of optional states. All pairwise comparisons are presented in Table 10.

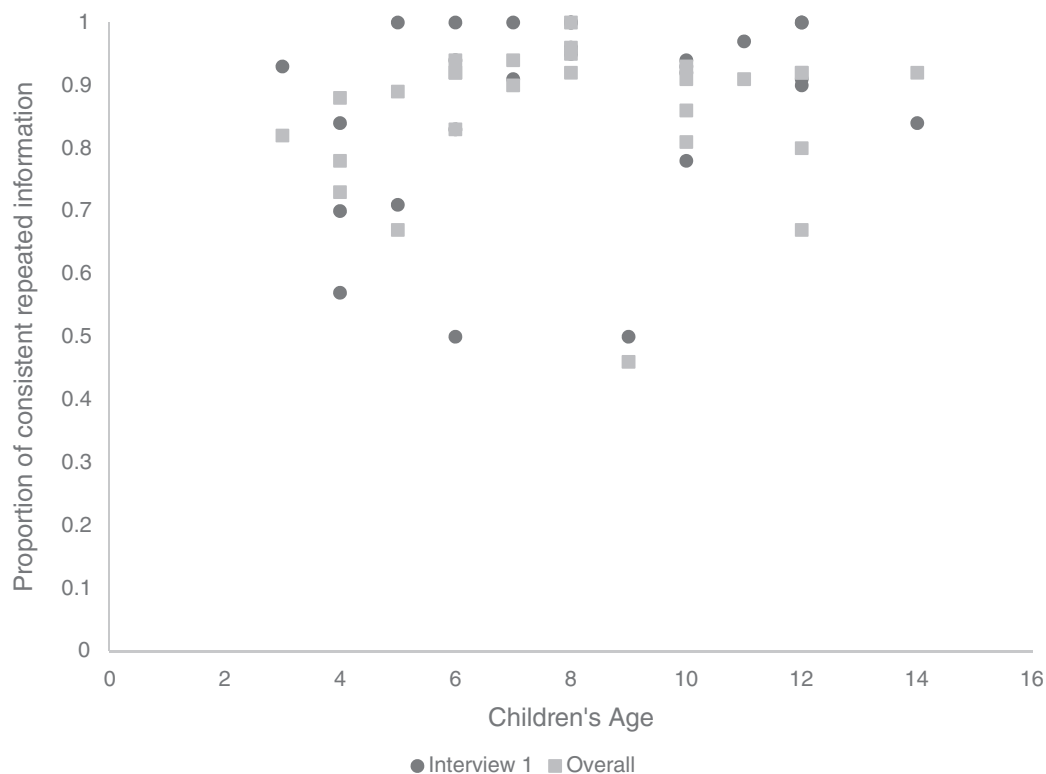
## 4 | DISCUSSION

Consistent with the hypotheses, results of the present study indicated that the completeness of children's testimonies increased significantly when information was collected over two interviews rather than a single interview. Not only did children report more new details over the course of two interviews, they also mentioned a higher number of crucial time and location details as well as subjective details related to

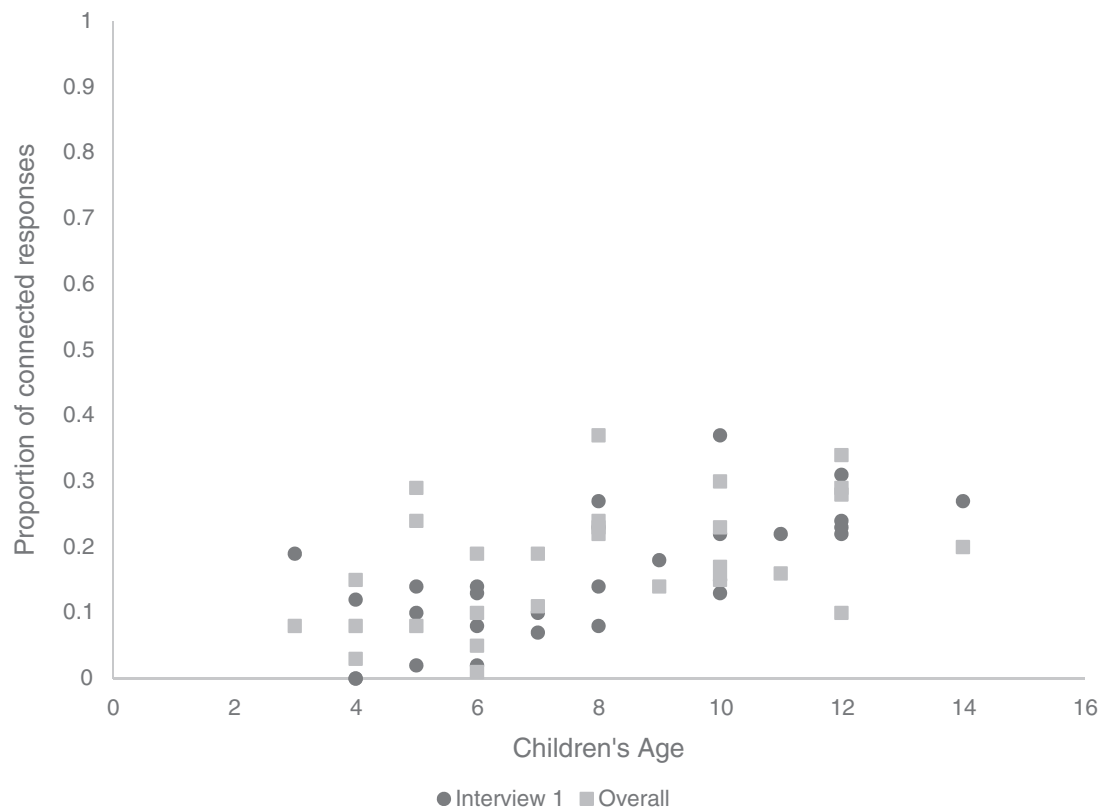
emotions, cognitions and perceptions, and sensitive details which were directly related to the alleged abuse. In fact, children reported a higher proportion of forensically relevant details in the second interview than in the first interview. Contrary to expectations, when both contradictions within the same interview and contradictions across interviews were considered, there was no significant difference in the proportion of contradictory repeated details between the first and second interviews. Exploratory analyses of linguistic connectedness revealed that the proportion of markers of connectedness remained stable across interviews. Children's age influenced the completeness and connectedness of their testimonies, and the impact of multiple interviews on the completeness of narratives.

**TABLE 9** Percentage of responses including linguistic markers of connectedness in Interview 1 and Interview 2

	Interview 1		Interview 2	
	M (%)	SD (%)	M (%)	SD (%)
Markers of connectedness	15.6	9	17.9	9
Simple temporal markers	3.79	5	4.51	6
Complex temporal markers	3.33	3	4.92	4
Causal relations	5.3	4	6.22	6
Optional states	3.2	3	2.26	3



**FIGURE 2** Children's age and the percentage of consistent information in Interview 1 and overall



**FIGURE 3** Children's age and the percentage of connected responses in Interview 1 and overall

**TABLE 10** Summary of analyses of the percentage of linguistic markers in children's responses in Interview 1 and overall

Analysis			Results			
Type	Variables	Df	Test value	p	Effect size	95% CI
Main effect	Connectedness	3,78	$F = 4.26$	$=.25$	$\eta_p^2 = 0.05$	
Main effect	Age	1,26	$F = 25.02$	$<.001^*$	$\eta_p^2 = 0.49$	
Main effect	Interview	1,26	$F = 1.41$	$=.008^*$	$\eta_p^2 = 0.14$	
Interaction	Connectedness X age	3,78	$F = 2.45$	$=.07$	$\eta_p^2 = 0.09$	
Interaction	Connectedness X interview	3,78	$F = 1.07$	$=.37$	$\eta_p^2 = 0.04$	
Interaction	Age X interview	1,26	$F = 1.30$	$=.27$	$\eta_p^2 = 0.06$	
Interaction	Age X interview X connectedness	3,78	$F = 0.62$	$=.60$	$\eta_p^2 = 0.02$	
Follow up	Simple tm vs. complex tm	26	$t = -0.29$	$=.78$	$d = 0.08$	$[-0.03, 0.02]$
Follow up	Simple tm vs. causal	26	$t = -1.08$	$=.29$	$d = 0.28$	$[-0.05, 0.02]$
Follow up	Simple tm vs. optional	26	$t = 1.86$	$=.08$	$d = 0.49$	$[-0.002, 0.05]$
Follow up	Complex tm vs. causal	26	$t = -1.08$	$=.29$	$d = 0.25$	$[-0.04, 0.01]$
Follow up	Complex tm vs. optional	26	$t = 2.74$	$=.01$	$d = 0.75$	$[0.007, 0.05]$
Follow up	Causal vs. optional	26	$t = 2.98$	$=.006^{**}$	$d = 0.83$	$[0.01, 0.07]$

Note: \*Denotes significance at  $p < .05$ , \*\* denotes significance at adjusted alpha level  $p < .008$ .

#### 4.1 | Completeness

Child witnesses in the current study provided more new information over the course of two interviews than in a single interview in every content category, except for information related to other witnesses.

Consistent with the results of Patterson and Pipe (2009), most details reported in the second interview were new rather than repeated. Crucially, interviewing children a second time also led to a significant increase in the number of time and location details available to investigators. Time and location details play a fundamental role in narrative

coherence; in multidimensional models, they provide the context of the narrative (Reese et al., 2011), whilst in story grammar approaches, they describe the temporal and physical setting for the story (Brown et al., 2018). Westcott and Kynan (2004) reported that in 26% of single forensic interviews analysed, children did not provide sufficient information about the physical setting of the events, and this number rose to 50% for temporal setting. Time and location details are vital for the particularisation of the alleged offenses, which is a prerequisite of the successful prosecution of child abuse cases (Powell & Thomson, 1997). Reflecting frequencies found in previous research (Connolly & Read, 2006), most child witnesses in this sample alleged that they have been victims of multiple instances of abuse. Particularisation is especially difficult when adults and children testify about multiple offences (e.g., Brubacher et al., 2014; Connolly & Gordon, 2014), as memories about repeated occurrences of an event are often organised into 'scripts' of general features and selecting a specific occurrence from several repetitions poses a source monitoring challenge (Johnson, Hashtroudi, & Lindsay, 1993). Therefore, the increase in the number of time and location details resulting from a second interview may have important implications for developing techniques to aid the process of particularisation when interviewing child witnesses alleging multiple abuse.

In addition to time and location details, subjective details relating to children's emotions, cognitive states and physical perceptions also increased in number over the course of two interviews. In the story-telling framework of legal decision making, first-person descriptions are essential components of credible testimonies (Pennington & Hastie, 1992). Subjective descriptions are also required for coherent narratives in the multidimensional model (Peterson et al., 2014; Reese et al., 2011) and in story grammar approaches (Brown et al., 2018). In the multidimensional model, subjective first-person descriptions are aspects of the 'evaluative' or 'emotional' dimension (Peterson et al., 2014; Reese et al., 2011), while in story grammar approaches, they are divided into the categories of 'internal response', referring to the emotions, cognitions and goals of the characters before the event and 'reaction', describing the emotions, cognitions and goals of the characters in response to the event (Brown et al., 2018). Previous field studies have shown that children only infrequently describe subjective reactions when describing sexual or physical abuse (Lyon, Scurich, Choi, Handmaker, & Blank, 2012; Westcott & Kynan, 2004), despite their ability to verbalise emotions in laboratory settings (Ahern & Lyon, 2013). The results of the present study indicate that in some contexts, multiple interviews can increase the number of subjective reactions children describe, including references to the emotional states, cognitions and perceptions of themselves and others.

Not only did child witnesses report a higher number of sensitive details overall than in a single interview, they also referred to a higher proportion of details directly related to the alleged offence in the second interview than in the first one. This tendency could potentially relate to children's increased comfort with the interview situation and increased trust in the interviewer during the second interview compared with the first one. This explanation is consistent with recommendations of the Extended Forensic Interview guide that

interviewers only approach the topic of the abuse in the second or third interview when talking to children alleging sexual abuse to allow sufficient rapport building (Carnes, Wilson, & Nelson-Gardell, 1999). Alternatively, the increased proportion of sensitive details might reflect children's increased understanding of the format of forensic interviews and the type of information interviewers are interested in. The way witnesses are expected to recall their memories is highly unusual for children, who are rarely asked to describe past events to this level of detail, especially to listeners who themselves are not knowledgeable about the event. Due to the question-answer format of forensic interviews, skilled interviewers can guide the conversation towards crucial topics even when using exclusively open-ended questions and carefully chosen follow-up questions (Lamb et al., 2007). In addition, previous research suggests that interviewers may ask more sensitive questions in further interviews, shifting from contextual to abuse-related details gradually over the course of multiple interviews (Patterson & Pipe, 2009). Whether due to better rapport between the child and the interviewer or to children's increased understanding of forensic interviews, the increased proportion of sensitive details in the second interview suggests that interviewers may gain a large amount of abuse-related information in multiple interviews which would not come to surface in single interviews.

## 4.2 | Consistency and connectedness

Based on the low consistency found in multiple interviews with children in previous research (Baugerud et al., 2014; Price et al., 2016), contradictions in children's recall were expected to increase in frequency across the two interview occasions. In contrast, when both contradictions within the same interview and across the two interviews were taken into account, results showed no significant difference in the proportion of consistent repeated details between the first and second interview. In line with previous findings (Katz & Hershkowitz, 2012; Waterhouse et al., 2016), there were few contradictions between interviews. The frequency of within-interview contradictions was similar in the two interviews and higher than the frequency of contradictions between interviews. Thus, although the overall number of contradictions increased in multiple interviews, the proportion of contradictions remained constant.

Research in forensic psychology cautions against the use of inconsistencies as correlates for accuracy (Gilbert & Fisher, 2006) but pointing out inconsistencies in witnesses' accounts is a common credibility challenging strategy during cross-examination (Szojka et al., 2017). The present findings suggest that multiple interviews do not necessarily damage the credibility of children's accounts by increasing the proportion of contradictions. However, inconsistencies are also introduced to children's accounts via omissions of previously mentioned details and additions of new details, and the proportion of these types of inconsistencies was very high in the current sample. Children's credibility may also be challenged on the basis of details that were 'left out' from earlier interviews and emerged in later interviews only (Szojka et al., 2017), but this is an inherent result of multiple



interviewing – additional interviews are conducted to reveal new details, but new details always mean more inconsistencies.

As no previous research investigated the extent to which forensic interviews with children elicit linguistically connected accounts, hypotheses regarding this component of narrative coherence were exploratory. In the present study, the proportion of linguistic connections in children's accounts remained stable across interviews. This result is consistent with those reported by Habermas and de Silveira (2008) who found that repeated recall opportunities did not affect linguistic connectedness in children's autobiographical accounts.

### 4.3 | Age and narrative coherence

The present study provided an insight into developmental differences in children's ability to provide coherent testimonies in response to multiple forensic interviews. Consistent with other studies, older children provided more complete responses in both interviews than younger children. Additionally, the linguistic connectedness of testimonies also increased as the age of witnesses increased. The age differences in the narrative coherence of children's testimony found in the present study are consistent with findings from research using a multidimensional model of narrative coherence indicating that thematic, chronological, contextual and evaluative elements of coherence develop at different rates across the lifespan (Habermas & de Silveira, 2008; Reese et al., 2011).

Although children of all ages provided more complete accounts over the course of the two interviews than in the first interview only, the difference in the number of details increased with children's age. These results are contrary to the increased benefit of multiple interviews for the youngest children reported by Baugerud et al. (2014) and Katz and Hershkowitz (2012). However, even though the increase in the number of details was proportionally higher for older children, a small increase can be valuable in forensic investigations where younger children often provide only brief descriptions in a single interview (Faller, Cordisco-Steele, & Nelson-Gardell, 2010; Langballe & Davik, 2017).

### 4.4 | Limitations and future research

Whilst the current research provided a crucial insight into the impact of multiple interviews on the narrative coherence of children's forensic testimonies, the design of the study left several questions open for future research. First, supporting research showing low adherence to best practice forensic interviewing guidelines (Lamb, 2016), few details in the present study were elicited by invitations and a substantial minority of details were provided in response to suggestive questions. Consistent with previous research (Waterhouse et al., 2016), there was no difference between the first interview and the second interview in the proportion of details elicited by any of the question types. However, in the second interview, invitations were less frequent than option-posing questions. This was not the case in the first interview, suggesting that interviewers may rely more on closed-ended questions and put less

emphasis on free recall in the second interview. As children's free recall is generally more accurate than their response to closed-ended questions (Lamb et al., 2007), this may raise concerns about the accuracy of children's recall in the second interview.

However, the field research design of the current study prevented analyses involving accuracy, as the 'ground truth' regarding the events children described was not known to the researchers. The impact of multiple interviews on the accuracy of children's accounts is a much-debated topic and research in the laboratory needs to explore the relationship between accuracy and narrative coherence further before any recommendations can be made about using multiple interviews as a method to scaffold the narrative coherence of child witnesses' accounts. Gaining complete, coherent and consistent accounts is essential for investigations, but inaccurate details can lead to miscarriages of justice (Ceci & Bruck, 1993).

Second, the sampling method of this study did not allow for the control of age and other witness characteristics. Due to the small number of witnesses and the large age differences in the current sample, it was not possible to investigate the relationship between age and more complex measures of narrative coherence, such as the content of children's responses. A larger sample size would allow researchers to categorise witnesses into meaningful age groups and offer more nuanced conclusions about the impact of multiple interviews in different age groups. In addition to age, there was also a large variability in other case characteristics, including the type of abuse, the interview protocol and the identity of the interviewer, thus, it is not possible to draw general conclusions from the present data.

Third, the type and quantity of complex real-life data provided in multiple forensic interviews prevented the use of coding schemes developed for multidimensional models of coherence (Reese et al., 2011), and local, quantitative measures of narrative coherence were used instead. Further research is needed to investigate the relationship between different models and coding approaches of narrative coherence and to determine whether legal and psychological models of coherence can be integrated.

In addition to exploring the relationship between narrative coherence and accuracy in multiple interviews, future research should clarify the impact of the three components of narrative coherence on the credibility of children's testimony. The story-telling model of legal decision making suggests that completeness, consistency and connectedness contribute independently to the credibility of witnesses' accounts (McAdams, 2006; Pennington & Hastie, 1988; Rideout, 2008). Although previous studies have reported that adults judge narratively coherent accounts as 'better stories' (Schneider & Winship, 2002), research has not yet investigated the effect of narrative coherence on mock jurors' credibility judgements of children's reports elicited in multiple interviews.

### 4.5 | Conclusion and implications

The present study was the first to investigate the narrative coherence of children's accounts elicited in multiple forensic interviews. Results

suggest that multiple interviews increase children's ability to provide coherent testimonies, through providing children with an opportunity to report more details related to crucial aspects of the allegations. Multiple interviews differentially affected the components of narrative coherence; while the completeness of children's accounts increased, their consistency and connectedness remained stable.

These results imply that investigators may conduct an additional forensic interview to increase the completeness of children's testimony, particularly if it lacks forensically relevant details related to the location and timing of the events and children's emotions, perceptions and cognitive appraisals. The findings also indicate that multiple interviews do not inherently increase the proportion of contradictions in children's accounts, suggesting that concerns about the potential negative effect of additional interviews on children's credibility may not be justified. However, as 'ground truth' is not known in field studies, the results of the present study should be interpreted in the context of laboratory research showing that the accuracy of children's recall across multiple interviews is contingent upon interview quality. Further research is needed to establish the relationship between narrative coherence and the accuracy and credibility of children's testimonies.

## CONFLICT OF INTEREST

The authors declare no conflict of interest in relation to this manuscript.

## DATA AVAILABILITY STATEMENT

Data sharing not applicable due to confidentiality and ethical considerations.

## ORCID

Zsolt A. Szojka  <https://orcid.org/0000-0002-4126-1261>

## REFERENCES

- Ahern, E. C., & Lyon, T. D. (2013). Facilitating maltreated children's use of emotional language. *Journal of Forensic Social Work*, 3(2), 176–203. <https://doi.org/10.1080/1936928X.2013.854124>
- Baugerud, G. A., Magnussen, S., & Melinder, A. (2014). High accuracy but low consistency in children's long-term recall of a real-life stressful event. *Journal of Experimental Child Psychology*, 126, 357–368. <https://doi.org/10.1016/j.jecp.2014.05.009>
- Bennett, W. L., & Feldman, M. S. (1981). *Reconstructing reality in the courtroom*. London: Tavistock Publications.
- Block, S. D., Foster, E. M., Pierce, M. W., Berkoff, M. C., & Runyan, D. K. (2013). Multiple forensic interviews during investigations of child sexual abuse: A cost-effectiveness analysis. *Applied Developmental Science*, 17(4), 174–183. <https://doi.org/10.1080/10888691.2013.836033>
- Brown, D. A., Brown, E.-J., Lewis, C. N., & Lamb, M. E. (2018). Narrative skill and testimonial accuracy in typically developing children and those with intellectual disabilities. *Applied Cognitive Psychology*, 32(5), 550–560. <https://doi.org/10.1002/acp.3427>
- Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2014). Recommendations for interviewing children about repeated experiences. *Psychology, Public Policy, and Law*, 20(3), 325–335. <https://doi.org/10.1037/law0000011>
- Carnes, C. N., Wilson, C., & Nelson-Gardell, D. (1999). Extended forensic evaluation when sexual abuse is suspected: A model and preliminary data. *Child Maltreatment*, 4(3), 242–254. <https://doi.org/10.1177/1077559599004003005>
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113(3), 403–439. <https://doi.org/10.1037/0033-2909.113.3.403>
- Connolly, D. A., & Gordon, H. M. (2014). Can order of general and specific memory prompts help children to recall an instance of a repeated event that was different from the others? *Psychology, Crime & Law*, 20(9), 852–864. <https://doi.org/10.1080/1068316X.2014.885969>
- Connolly, D. A., & Read, J. (2006). Delayed prosecutions of historic child sexual abuse: Analyses of 2064 Canadian criminal complaints. *Law and Human Behavior*, 30(4), 409–434. <https://doi.org/10.1007/s10979-006-9011-6>
- Erdelyi, M. H. (1996). *The recovery of unconscious memories: Hypernesia and reminiscence*. University of Chicago Press.
- Faller, K. C., Cordisco-Steele, L., & Nelson-Gardell, D. (2010). Allegations of sexual abuse of a child: What to do when a single forensic interview isn't enough. *Journal of Child Sexual Abuse*, 19(5), 572–589. <https://doi.org/10.1080/10538712.2010.511985>
- Feltis, B. B., Powell, M. B., & Roberts, K. P. (2011). The effect of event repetition on the production of story grammar in children's event narratives. *Child Abuse & Neglect*, 35(3), 180–187. <https://doi.org/10.1016/j.chiabu.2010.11.004>
- Feltis, B. B., Powell, M. B., Snow, P. C., & Hughes-Scholes, C. H. (2010). An examination of the association between interviewer question type and story-grammar detail in child witness interviews about abuse. *Child Abuse & Neglect*, 34(6), 407–413. <https://doi.org/10.1016/j.chiabu.2009.09.019>
- Fogliati, R., & Bussey, K. (2014). The effects of cross-examination on children's reports of neutral and transgressive events. *Legal and Criminological Psychology*, 19(2), 296–315. <https://doi.org/10.1111/lcrp.12010>
- Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology*, 20(6), 723–739. <https://doi.org/10.1002/acp.1232>
- Goodman, G. S., & Quas, J. A. (2008). Repeated interviews and children's memory: It's more than just how many. *Current Directions in Psychological Science*, 17(6), 386–390. <https://doi.org/10.1111/j.1467-8721.2008.00611.x>
- Graffam Walker, A., Kenniston, J., Inada, S. S., & Caldwell, C. (2013). *Handbook on questioning children: A linguistic perspective*. ABA Center on Children and the Law. <https://www.americanbar.org/products/inv/book/213559/>
- Habermas, T., & Bluck, S. (2000). Getting a life: The emergence of the life story in adolescence. *Psychological Bulletin*, 126(5), 748–769. <https://doi.org/10.1037/0033-2909.126.5.748>
- Habermas, T., & de Silveira, C. (2008). The development of global coherence in life narratives across adolescence: Temporal, causal, and thematic aspects. *Developmental Psychology*, 44(3), 707–721. <https://doi.org/10.1037/0012-1649.44.3.707>
- Hershkovitz, I., & Terner, A. (2007). The effects of repeated interviewing on children's forensic statements of sexual abuse. *Applied Cognitive Psychology*, 21(9), 1131–1143. <https://doi.org/10.1002/acp.1319>
- Jack, F., & Zajac, R. (2014). The effect of age and reminders on witnesses' responses to cross-examination-style questioning. *Journal of Applied Research in Memory and Cognition*, 3(1), 1–6. <https://doi.org/10.1016/j.jarmac.2013.12.001>
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3–28. <https://doi.org/10.1037/0033-2909.114.1.3>
- Katz, C., & Hershkovitz, I. (2012). Repeated Interviews With children who are the alleged victims of sexual abuse. *Research on Social Work Practice*, 23(2), 210–218.
- Knutsson, J., Allwood, C. M., & Johansson, M. (2011). Child and adult witnesses: The effect of repetition and invitation-probes on free recall and metamemory realism. *Metacognition and Learning*, 6(3), 213–228. <https://doi.org/10.1007/s11409-011-9071-y>
- Krix, A. C., Sauerland, M., Lorei, C., & Rispens, I. (2015). Consistency across repeated eyewitness interviews: Contrasting police detectives' beliefs

- with actual eyewitness performance. *PloS one*, 10(2). <https://doi.org/10.1371/journal.pone.0118641>
- Kulkofsky, S., Wang, Q., & Ceci, S. J. (2008). Do better stories make better memories? Narrative quality and memory accuracy in preschool children. *Applied Cognitive Psychology*, 22(1), 21–38. <https://doi.org/10.1002/acp.1326>
- Lamb, M. E. (2016). Difficulties translating research on forensic interview practices to practitioners: Finding water, leading horses, but can we get them to drink? *American Psychologist*, 71(8), 710. <https://doi.org/10.1037/amp0000039>
- La Rooy, D., Lamb, M. E., & Pipe, M.-E. (2009). Repeated interviewing: A critical evaluation of the risks and potential benefits. In Kuehnle, K., & Connell, M. (Eds.). *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 327–361). Hoboken, NJ: Wiley.
- La Rooy, D. L., Pipe, M.-E., & Murray, J. E. (2005). Reminiscence and hypernesia in children's eyewitness memory. *Journal of Experimental Child Psychology*, 90(3), 235–254. <https://doi.org/10.1016/j.jecp.2004.11.002>
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). Structured forensic interview protocols improve the quality and informativeness of investigative interviews with children: A review of research using the NICHD investigative interview protocol. *Child Abuse & Neglect*, 31(11–12), 1201–1231. <https://doi.org/10.1016/j.chiabu.2007.03.021>
- Langballe, Å., & Davik, T. (2017). Sequential interviews with preschool children in Norwegian Barnahus. In S. Johansson, K. Stefansen, E. Bakkeiteig, & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic Barnahus model* (pp. 165–183). Palgrave macmillan. [https://doi.org/10.1007/978-3-319-58388-4\\_8](https://doi.org/10.1007/978-3-319-58388-4_8)
- Leander, L. (2010). Police interviews with child sexual abuse victims: Patterns of reporting, avoidance and denial. *Child Abuse & Neglect*, 34(3), 192–205. <https://doi.org/10.1016/j.chiabu.2009.09.011>
- London, K., Bruck, M., & Melnyk, L. (2009). Post-event information affects children's autobiographical memory after one year. *Law and Human Behavior*, 33(4), 344–355. <https://doi.org/10.1007/s10979-008-9147-7>
- Lyon, T. D., Scurich, N., Choi, K., Handmaker, S., & Blank, R. (2012). "How did you feel?": Increasing child sexual abuse witnesses' production of evaluative information. *Law and Human Behavior*, 36(5), 448–457. <https://doi.org/10.1037/h0093986>
- McAdams, D. P. (2006). The problem of narrative coherence. *Journal of Constructivist Psychology*, 19(2), 109–125. <https://doi.org/10.1080/10720530500508720>
- Melinder, A., Alexander, K., Cho, Y. I., Goodman, G. S., Thoresen, C., Lonnum, K., & Magnussen, S. (2010). Children's eyewitness memory: A comparison of two interviewing strategies as realized by forensic professionals. *Journal of Experimental Child Psychology*, 105(3), 156–177. <https://doi.org/10.1016/j.jecp.2009.04.004>
- Morris, G., Baker-Ward, L., & Bauer, P. J. (2010). What remains of that day: The survival of children's autobiographical memories across time. *Applied Cognitive Psychology*, 24(4), 527–544. <https://doi.org/10.1002/acp.1567>
- Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, 111(2), 486–511. <https://doi.org/10.1037/0033-295X.111.2.486>
- O'Neill, S., & Zajac, R. (2013). The role of repeated interviewing in children's responses to cross-examination-style questioning: Cross-examination and repeated interviewing. *British Journal of Psychology*, 104(1), 14–38. <https://doi.org/10.1111/j.2044-8295.2011.02096.x>
- Otgaar, H., Verschuere, B., Meijer, E. H., & van Oorsouw, K. (2012). The origin of children's implanted false memories: Memory traces or compliance? *Acta Psychologica*, 139(3), 397–403. <https://doi.org/10.1016/j.actpsy.2012.01.002>
- Patterson, T., & Pipe, M.-E. (2009). Exploratory assessments of child abuse: Children's responses to interviewer's questions across multiple interview sessions. *Child Abuse & Neglect*, 33(8), 490–504. <https://doi.org/10.1016/j.chiabu.2008.12.012>
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 521–533. <https://doi.org/10.1037/0278-7393.14.3.521>
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, 62(2), 189–206. <https://doi.org/10.1037/0022-3514.62.2.189>
- Peterson, C. (2011). Children's memory reports over time: Getting both better and worse. *Journal of Experimental Child Psychology*, 109(3), 275–293. <https://doi.org/10.1016/j.jecp.2011.01.009>
- Peterson, C. (2015). A decade later: Adolescents' memory for medical emergencies: A decade later. *Applied Cognitive Psychology*, 29(6), 826–834. <https://doi.org/10.1002/acp.3192>
- Peterson, C., Morris, G., Baker-Ward, L., & Flynn, S. (2014). Predicting which childhood memories persist: Contributions of memory characteristics. *Developmental Psychology*, 50(2), 439–448. <https://doi.org/10.1037/a0033221>
- Powell, M. B., & Thomson, D. M. (1997). The effect of an intervening interview on children's ability to remember one occurrence of a repeated event. *Legal and Criminological Psychology*, 2(2), 247–262. <https://doi.org/10.1111/j.2044-8333.1997.tb00346.x>
- Price, H. L., Connolly, D. A., & Gordon, H. M. (2016). Children who experienced a repeated event only appear less accurate in a second interview than those who experienced a unique event. *Law and Human Behavior*, 40(4), 362–373. <https://doi.org/10.1037/lhb0000194>
- Reese, E., Haden, C. A., Baker-Ward, L., Bauer, P., Fivush, R., & Ornstein, P. A. (2011). Coherence of personal narratives across the lifespan: A multidimensional model and coding method. *Journal of Cognition and Development*, 12(4), 424–462. <https://doi.org/10.1080/15248372.2011.587854>
- Rideout, J. C. (2008). Storytelling, narrative rationality, and legal persuasion. *Legal Writing: The Journal of the Legal Writing Institute*, 14, 53.
- Righarts, S., Jack, F., Zajac, R., & Hayne, H. (2015). Young children's responses to cross-examination style questioning: The effects of delay and subsequent questioning. *Psychology, Crime & Law*, 21(3), 274–296. <https://doi.org/10.1080/1068316X.2014.951650>
- Schneider, B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-subject factors. *Frontiers in Psychology*, 6, 474. <https://doi.org/10.3389/fpsyg.2015.00474>
- Schneider, P., & Winship, S. (2002). Adults' judgments of fictional story quality. *Journal of Speech, Language, and Hearing Research: JSLHR*, 45(2), 372–383. [https://doi.org/10.1044/1092-4388\(2002\)029](https://doi.org/10.1044/1092-4388(2002)029)
- Szojka, Z. A., Andrews, S. J., Lamb, M. E., Stolzenberg, S. N., & Lyon, T. D. (2017). Challenging the credibility of alleged victims of child sexual abuse in Scottish courts. *Psychology, Public Policy, and Law*, 23(2), 200. <https://doi.org/10.1037/law0000117>
- Waterhouse, G. F., Ridley, A. M., Bull, R., La Rooy, D., & Wilcock, R. (2016). Dynamics of repeated interviews with children: Repeated interviews with children. *Applied Cognitive Psychology*, 30(5), 713–721. <https://doi.org/10.1002/acp.3246>
- Westcott, H. L., & Kynan, S. (2004). The application of a 'story-telling' framework to investigative interviews for suspected child sexual abuse. *Legal and Criminological Psychology*, 9(1), 37–56. <https://doi.org/10.1348/135532504322776843>

**How to cite this article:** Szojka ZA, Nicol A, La Rooy D.

Narrative coherence in multiple forensic interviews with child witnesses alleging physical and sexual abuse. *Appl Cognit Psychol*. 2020;1–18. <https://doi.org/10.1002/acp.3673>

## APPENDIX

As the probability of Type 1 errors is elevated in within-subject designs involving an interaction between the covariate and the dependent variable (Schneider et al., 2015), alternative statistical tests were conducted corresponding to each ANCOVA, without the covariate.

### Number of details

To assess whether children reported significantly more details overall in the two interviews than in the first one only, a paired *t* test was conducted comparing the number of new details children recalled in the first interview and across the two interviews. Significantly less new details were recalled in a single interview ( $M = 121.75$ ,  $SD = 147.48$ ) than across the two interviews ( $M = 221.04$ ,  $SD = 250.70$ ),  $t(27) = -4.27$ ,  $p < .001$ ,  $d = 0.48$ , 95% CI =  $[-147.03, -51.54]$ .

### Consistency

To assess whether the proportion of consistent responses was affected by interview number, a paired *t* test was conducted

comparing the consistency of repeated information in the first interview and across the two interviews. There was no significant difference between the proportion of consistent responses in the first interview ( $M = 0.87$ ,  $SD = 0.15$ ) and overall ( $M = 0.84$ ,  $SD = 0.16$ ),  $t(27) = 0.61$ ,  $p = .55$ ,  $d = 0.08$ , 95% CI =  $[-0.03, 0.06]$ .

### Connectedness

To assess whether children's accounts were significantly more connected in the second interview than in the first one, a RM-ANOVA was conducted assessing the potential effect of interview number (first interview, overall) on the frequency of the four types of connectedness markers (simple temporal, complex temporal, causal, optional) witnesses reported. A significant main effect was found for connectedness,  $F(3, 81) = 4.04$ ,  $p = .01$ ,  $\eta_p^2 = 0.13$ . No main effect was found for interview number,  $F(1, 27) = 1.39$ ,  $p = .25$ ,  $\eta_p^2 = 0.05$ . No significant interactions were found.